Predictive abilities of the Unpaired and Paired Release Models
Ken B. Newman, January 24, 2005

This brief note summarizes a comparison between the unpaired release model (Newman and Rice 2002[1]) and the paired release model (Newman 2003[2]) for recoveries of experimental fall chinook salmon smolt releases made in the lower Sacramento River. The comparison is made solely with regard to the two models' ability to predict recoveries for a "test data set", namely recoveries at Chipps Island of releases that were not used to fit the two models.

Test data set

The release and recovery numbers and covariates for the test data set are shown in Table 2. There are a total of 50 release groups shown in the table but only 40 have complete covariate data. Missing covariate values were "imputed" in an *ad hoc* manner for three of the releases. In particular, the tide variable values were set equal to zero for the fifth and sixth release groups (Tag codes: 06-01-06-02-11 and 06-01-06-02-12); the coefficient for tide is quite small and insignificant in both models and the effect of inserting nearly any value for tide should be quite small. The other imputation was for the 38th release (Tag code: 06-27-17) which had zero recoveries and therefore no flow, nor salinity, information; the flow and salinity values for a 'replicate' release (37th release, Tag code: 06-27-16) were used. Unfortunately for six of the release groups (the last six in the table) there was no turbidity information, and the coefficient for turbidity is relatively large, and no related releases existed that could supply surrogate information.

Note: more formal imputation procedures are possible, but take more work. For example, the missing turbidity values could be estimated by using regression model predictions of turbidity as a function of the other covariates, where the regression model is fit using observations with complete data.

Methods

Predictions using the unpaired release model, i.e., the generalized linear model with ridge parameter, were based on the coefficients shown in the column labelled β_λ in Table 1 of Newman and Rice (2002). The particular version of the unpaired release model used was the hierarchical model with capture probabilities fixed and a shock effect allowed for downstream releases (the mean values of the coefficients are shown in the last column of Table 5 of Newman (2003)).

In each case the covariates were first standardized, using means and standard deviations based upon the original data sets used to fit the models. For the unpaired model, just point estimates of the number of recoveries were calculated, with estimates made as follows:

$$\hat{y}_r \quad = \quad R \times f \times \exp(\beta_\lambda' \mathbf{x}), \tag{1}$$

where R is the number released and $\mathbf{x}$ are the standardized covariate values. The coefficient f

[1]Newman, K.B., and Rice, J. 2002. "Modeling the survival of chinook salmon smolts outmigrating through the lower Sacramento river system." *Journal of the American Statistical Association* **97**: 983–993.

[2]Newman, K.B. 2003. "Modelling paired release-recovery data in the presence of survival and capture heterogeneity with application to marked juvenile salmon." *Statistical Modelling* **3**: 157–177.

represents the Chipps Island trawl net efficiency; the value 0.11 was used and was the median value shown in Newman and Rice (1997[3]).

For the paired model, a large sample (48,000) of values were drawn from the posterior distribution of the coefficients. For each sampled value (a vector of coefficients), the survival and capture probabilities were generated from logit-normal distributions; i.e.,

$$\begin{aligned} logit(S) &= \beta' \mathbf{x1} + \epsilon_s \quad \text{where } \epsilon_s \sim \text{Normal}(0, \sigma_s^2) && (2) \\ logit(p) &= \delta' \mathbf{x2} + \epsilon_p \quad \text{where } \epsilon_p \sim \text{Normal}(0, \sigma_p^2) && (3) \end{aligned}$$

where β and δ are the vector of coefficients for survival and capture rates (with $x1$ and $x2$ being the corresponding covariates) and σ_s^2 and σ_p^2 are the variances of the random effects. (For the $x2$ values only the intercept was used; i.e., the non-1988 capture levels were assumed—this just affects scaling of the predictions) The survival and capture rates were then calculated by the inverse logit transform, i.e.,

$$\begin{aligned} S &= \exp\left(logit(S)\right) / \left(1 + \exp(logit(S))\right) && (4) \\ p &= \exp\left(logit(p)\right) / \left(1 + \exp(logit(p))\right). && (5) \end{aligned}$$

Then the number of recoveries was estimated as follows:

$$\hat{y} = R \times S \times p. \quad (6)$$

Finally, the mean of the resulting 48,000 estimates was also calculated.

Note that one distinction between the covariates used for the two models is that export to inflow ratio is used for the unpaired model, while the paired model uses absolute exports.

Results

Table 1 compares the predicted numbers (and percentage recovery rates) for the unpaired and paired models. In terms of median absolute errors, $|Obs - Pred|$, the unpaired release model had a value of 9.5 versus 13.7 for the paired release model.

Scatterplots of the observed recoveries against the predicted recoveries for the two models are shown in the top plots of Figure 1 along with linear regression lines. In terms of the the degree of correlation between predictions and observations, the paired release model is slightly better with a Pearson correlation coeffient, r, of 0.71, versus an r of 0.60 for the unpaired model.

Another way to compare the models is to examine relative recovery rates for different releases (possibly under different environmental conditions). If capture rates are the same for both releases, then relative recovery rates are the same as relative survival rates. The observed and predicted recovery rates, number recovered/R, for all 43 releases were calculated. The highest observed recovery rate (Tag Code: 06-01-06-05-07, which had 32 recoveries from 21,380 release) was then used as a benchmark for comparing recovery rates for other releases:

$$\text{Relative recovery rate of best to release X} = \frac{\text{Highest Rec Rate}}{\text{Recovery rate of release X}} = \frac{y/R\ [best]}{y/R\ [X]}.$$

[3]Newman, K.B., and Rice, J. 1997 "Statistical Model for Survival of Chinook Salmon Smolts Outmigrating through the Lower Sacramento-San Joaquin System." Interagency Ecological Program, Technical Report 59.

Table 1: Predicted recoveries (in numbers and percentages relative to number released) based on paired and unpaired models.

	Tag Codes	Release site (stock)	Number Released	Obs'd No.	Unpaired No.	Paired No.	Observed %	Unpaired %	Paired %
1	06-01-14-05-02	Ryde (FRH)	51597	47	88	127	0.091	0.170	0.247
2	06-01-14-05-03	Miller Park (FRH)	50292	32	75	119	0.064	0.150	0.237
3	06-01-06-01-08	Miller Park (FRH)	49708	58	63	92	0.117	0.127	0.184
4	06-01-06-02-02	Miller Park (FRH)	49881	30	32	45	0.060	0.065	0.090
5	06-01-06-02-11	West Sacramento (FRH)	25642	14	26	23	0.055	0.102	0.089
6	06-01-06-02-12	West Sacramento (FRH)	25032	9	25	21	0.036	0.100	0.085
7	06-01-06-02-13	West Sacramento (FRH)	25829	15	20	17	0.058	0.077	0.068
8	06-01-06-02-14	West Sacramento (FRH)	26315	7	24	18	0.027	0.092	0.069
9	06-01-06-02-10	West Sacramento (FRH)	25069	1	11	6	0.004	0.042	0.026
10	06-01-06-05-07	West Sacramento (FRH)	21380	32	37	49	0.150	0.173	0.229
11	06-01-06-05-08	West Sacramento (FRH)	21556	26	39	50	0.121	0.182	0.231
12	06-01-06-05-09	West Sacramento (FRH)	17830	37	30	41	0.208	0.170	0.229
13	06-01-06-05-10	West Sacramento (FRH)	16498	29	31	38	0.176	0.190	0.230
14	05-01-02-07-14	Ryde (FRH)	25873	13	64	56	0.050	0.249	0.217
15	05-01-02-07-15	Ryde (FRH)	25133	10	59	55	0.040	0.235	0.217
16	05-24-16	West Sacramento (FRH)	25621	28	55	52	0.109	0.213	0.202
17	05-24-17	West Sacramento (FRH)	26174	34	61	51	0.130	0.235	0.196
18	05-24-14	Ryde (FRH)	26489	34	43	42	0.128	0.161	0.160
19	05-24-15	Ryde (FRH)	25814	30	37	42	0.116	0.144	0.161
20	05-23-24	West Sacramento (FRH)	25695	20	27	34	0.078	0.104	0.132
21	05-23-25	West Sacramento (FRH)	25977	18	28	36	0.069	0.106	0.138
22	05-01-02-07-06	West Sacramento (FRH)	25585	10	21	24	0.039	0.080	0.093
23	05-01-02-07-07	West Sacramento (FRH)	25633	10	21	24	0.039	0.080	0.093
24	05-01-02-08-05	Ryde (FRH)	23042	5	33	29	0.022	0.142	0.128
25	05-01-02-08-06	Ryde (FRH)	23468	6	22	31	0.026	0.096	0.130
26	06-26-55	West Sacramento (FRH)	25005	15	17	31	0.060	0.068	0.124
27	06-26-56	West Sacramento (FRH)	25011	19	15	31	0.076	0.062	0.123
28	05-01-02-08-07	Ryde (FRH)	21419	9	15	29	0.042	0.072	0.135
29	05-01-02-08-08	Ryde (FRH)	21395	12	15	28	0.056	0.071	0.133
30	06-26-53	West Sacramento (FRH)	20926	21	23	25	0.100	0.108	0.121
31	06-26-54	West Sacramento (FRH)	20613	23	22	25	0.112	0.107	0.122
32	06-26-49	West Sacramento (FRH)	17416	19	18	18	0.109	0.104	0.105
33	06-26-50	West Sacramento (FRH)	17064	12	18	18	0.070	0.104	0.105
34	06-27-14	West Sacramento (MOK)	25795	24	17	19	0.093	0.067	0.073
35	06-27-15	West Sacramento (MOK)	25795	25	14	18	0.097	0.054	0.071
36	06-27-16	West Sacramento (MOK)	25757	1	14	15	0.004	0.054	0.058
37	06-27-17	West Sacramento (MOK)	25757	0	14	15	0.000	0.054	0.058
38	06-27-08	West Sacramento (MOK)	25872	6	7	7	0.023	0.029	0.025
39	06-27-09	West Sacramento (MOK)	25872	8	8	7	0.031	0.031	0.025
40	06-26-95	West Sacramento (FRH)	44563	19	24	28	0.043	0.053	0.063
41	06-26-99	Ryde (FRH)	43789	8	58	43	0.018	0.132	0.099
42	06-26-97	West Sacramento (FRH)	45972	27	54	47	0.059	0.118	0.102
43	06-26-98	Ryde (FRH)	43676	14	39	34	0.032	0.088	0.078

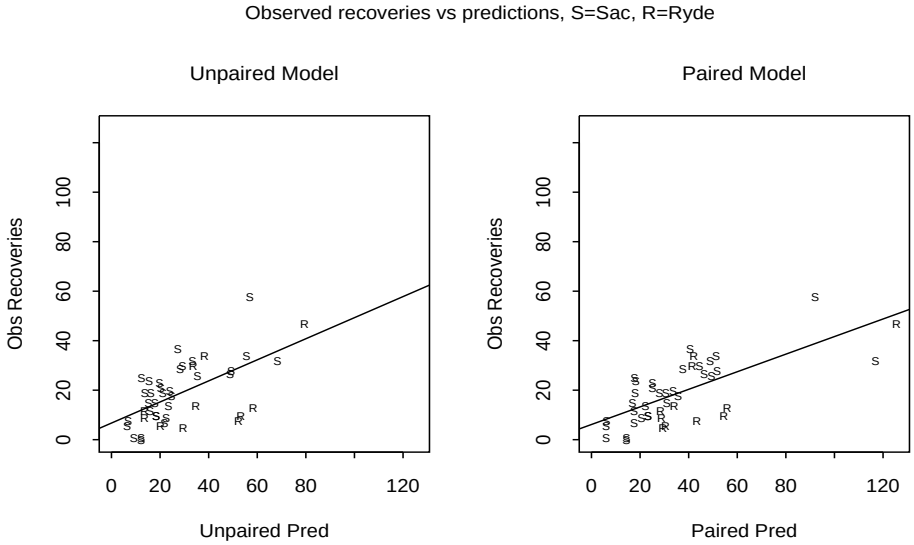
The ratios of predicted recovery rates (with Tag Code: 06-01-06-05-07 as the benchmark in all cases) were then calculated for both the unpaired and paired release models. The absolute differences between observed and predicted relative recovery rates were then calculated. The two models were virtually identical; see the density plots of the absolute differences in the bottom plot of Figure 1.

To make the comparison between the models a little more similar, the paired release model was fit using export to inflow ratio instead of total exports. The results were similar. The median absolute error for the paired release model decreased slightly to 12.7 and the correlation was r=0.68.

Conclusions

The unpaired release model, with trawl efficiency set at 0.11, produced predictions of numbers of recoveries that were closer to the observed recoveries than did the paired release model (with capture rate set at the non-1988 levels). For predicting the relative recovery rates of two different release groups (implicitly, perhaps, relative survival rates), the two models appear roughly equivalent.

Figure 1: Top plots are of the observed number of recoveries versus predicted values for both the unpaired and paired release models. Bottom plot contains density plots of the absolute errors in relative recovery rates for the two models (unpaired is black solid line; paired is green dashed line).



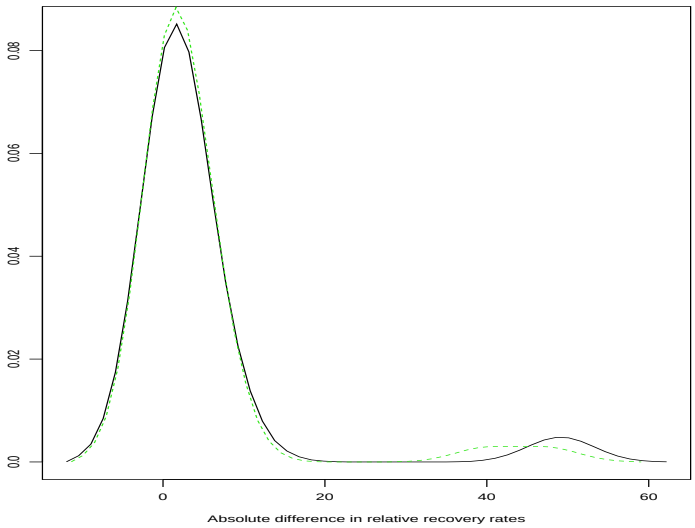


Table 2: Test data set.

Tag Code	Site(Stock)	H.Temp	Size	#Rel	#Rec	Rel.Temp	Truck	Shock	Log(Flow)	Salinity	Tide	Up.EI	Up.Gate	Turbid	Exports
06-01-14-05-02	Ryde (FRH)	48	81	51597	47	59.0	57.0	2.0	11.35	141.3	1.25	0.03	0	50.0	4093.5
06-01-14-05-03	Miller Park (FRH)	48	81	50292	32	59.0	60.0	-1.0	11.26	138.4	1.25	0.03	0	50.0	4093.5
06-01-06-01-08	Miller Park (FRH)	48	81	49708	58	57.0	48.0	9.0	10.17	146.9	1.96	0.05	0	6.0	1783.0
06-01-06-02-02	Miller Park (FRH)	52	83	49881	30	65.0	58.0	7.0	9.81	128.3	1.72	0.06	0	6.0	1772.0
06-01-06-02-11	West Sacramento (FRH)	49	95	25642	14	63.0	48.0	15.0	9.37	354.2	NA	0.11	0	7.5	2458.0
06-01-06-02-12	West Sacramento (FRH)	49	95	25032	9	63.0	48.0	15.0	9.31	360.8	NA	0.12	0	7.5	2474.0
06-01-06-02-13	West Sacramento (FRH)	51	102	25829	15	65.0	49.0	16.0	9.21	420.5	0.74	0.13	0	7.0	2540.0
06-01-06-02-14	West Sacramento (FRH)	51	102	26315	7	65.0	49.0	16.0	9.22	443.1	-0.16	0.13	0	7.0	2534.0
06-01-06-02-09	West Sacramento (FRH)	52	107	25152	0	72.0	52.0	20.0	NA	443.1	NA	0.13	0	7.0	2534.0
06-01-06-02-10	West Sacramento (FRH)	52	107	25069	1	72.0	52.0	20.0	9.25	1859.8	-0.08	0.12	1	8.5	2394.0
06-01-06-05-07	West Sacramento (FRH)	48	83	21380	32	56.0	50.0	6.0	10.77	148.1	1.64	0.03	0	13.5	2199.0
06-01-06-05-08	West Sacramento (FRH)	48	83	21556	26	56.0	50.0	6.0	10.77	117.8	1.29	0.03	0	17.0	2187.0
06-01-06-05-09	West Sacramento (FRH)	48	99	17830	37	57.0	54.0	3.0	10.82	208.3	2.57	0.05	0	10.0	3976.0
06-01-06-05-10	West Sacramento (FRH)	48	99	16498	29	57.0	54.0	3.0	10.82	209.0	1.97	0.05	0	10.0	3878.0
05-01-02-07-14	Ryde (FRH)	47	75	25873	13	53.0	54.0	-1.0	10.35	240.2	1.00	0.12	0	12.5	5164.0
05-01-02-07-15	Ryde (FRH)	47	75	25133	10	53.0	54.0	-1.0	10.34	252.3	1.27	0.12	0	15.0	5101.5
05-24-16	West Sacramento (FRH)	50	82	25621	28	55.0	59.0	-4.0	10.37	210.1	0.93	0.09	0	10.0	3449.0
05-24-17	West Sacramento (FRH)	50	82	26174	34	55.0	59.0	-4.0	10.28	211.5	0.32	0.09	0	10.0	3449.0
05-24-14	Ryde (FRH)	50	86	26489	34	61.0	55.0	6.0	10.10	198.6	-0.12	0.10	0	10.0	3478.0
05-24-15	Ryde (FRH)	50	86	25814	30	61.0	55.0	6.0	10.12	187.7	0.46	0.10	0	10.0	3477.0
05-23-24	West Sacramento (FRH)	50	82	25695	20	61.0	55.0	6.0	10.02	192.2	1.07	0.10	0	10.0	3497.0
05-23-25	West Sacramento (FRH)	50	82	25977	18	61.0	55.0	6.0	10.10	193.4	1.07	0.10	0	10.0	3497.0
05-01-02-07-06	West Sacramento (FRH)	50	87	25585	10	64.0	54.0	10.0	9.75	213.9	0.84	0.14	0	9.0	3546.5
05-01-02-07-07	West Sacramento (FRH)	50	87	25633	10	64.0	54.0	10.0	9.75	218.0	0.84	0.14	0	9.0	3537.0
05-01-02-08-05	Ryde (FRH)	48	71	23042	5	59.0	50.0	9.0	10.15	203.8	-0.10	0.28	0	5.0	8882.5
05-01-02-08-06	Ryde (FRH)	48	71	23468	6	59.0	50.0	9.0	10.18	202.9	2.05	0.27	0	5.0	8735.0
06-26-55	West Sacramento (FRH)	53	77	25005	15	63.0	56.0	7.0	10.23	161.1	2.35	0.07	0	7.0	2506.0
06-26-56	West Sacramento (FRH)	53	77	25011	19	63.0	56.0	7.0	10.21	161.9	2.87	0.07	0	7.0	2463.0
05-01-02-08-07	Ryde (FRH)	51	79	21419	9	64.0	56.0	8.0	10.15	156.1	2.41	0.07	0	7.0	2230.0
05-01-02-08-08	Ryde (FRH)	51	79	21395	12	64.0	56.0	8.0	10.13	156.9	2.41	0.07	0	7.0	2224.0
06-26-53	West Sacramento (FRH)	52	83	20926	21	64.0	57.0	7.0	10.13	170.7	-0.23	0.07	0	7.0	2224.0
06-26-54	West Sacramento (FRH)	52	83	20613	23	64.0	57.0	7.0	10.15	154.8	-0.10	0.06	0	7.0	2212.0
06-26-49	West Sacramento (FRH)	52	93	17416	19	63.0	54.0	9.0	9.75	200.2	0.42	0.11	0	7.0	2865.0
06-26-50	West Sacramento (FRH)	52	93	17064	12	63.0	54.0	9.0	9.75	189.4	0.42	0.11	0	7.0	2865.0
06-27-14	West Sacramento (MOK)	51	81	25795	24	64.0	55.0	9.0	9.36	1179.1	1.44	0.08	0	9.5	1554.0
06-27-15	West Sacramento (MOK)	51	81	25795	25	64.0	55.0	9.0	9.35	1040.8	2.57	0.08	0	9.5	1690.5
06-27-16	West Sacramento (MOK)	49	78	25757	1	66.0	54.0	12.0	9.31	1263.0	1.38	0.09	0	9.5	1893.5
06-27-17	West Sacramento (MOK)	49	78	25757	0	66.0	54.0	12.0	NA	1263.0	NA	0.09	0	9.5	1893.5
06-27-08	West Sacramento (MOK)	49	85	25872	6	72.0	54.0	18.0	8.92	2618.3	2.06	0.12	0	7.0	1834.0
06-27-09	West Sacramento (MOK)	49	85	25872	8	72.0	54.0	18.0	8.92	2777.8	1.77	0.12	0	7.0	1834.0
06-26-95	West Sacramento (FRH)	51	59	44563	19	64.0	50.0	14.0	9.64	490.3	1.12	0.11	0	4.6	2283.0
06-26-99	Ryde (FRH)	51	74	43789	8	60.8	51.8	9.0	9.65	518.6	-0.62	0.22	0	4.6	4707.5
06-26-97	West Sacramento (FRH)	49	89	45972	27	60.8	51.8	9.0	9.37	541.6	0.34	0.09	0	5.0	1661.0
06-26-98	Ryde (FRH)	52	83	43676	14	64.4	51.8	12.6	9.36	507.8	0.34	0.09	0	5.0	1664.0
06-01-08-08-04	West Sacramento (FRH)	51	72	26069	16	57.0	54.0	3.0	10.14	168.4	2.52	0.04	0	NA	1778.5
06-01-08-08-05	West Sacramento (FRH)	52	72	26096	28	57.0	54.0	3.0	10.02	170.6	2.52	0.05	0	NA	1709.0
06-01-08-08-00	West Sacramento (FRH)	52	79	25783	34	54.5	51.8	2.7	10.21	147.5	3.21	0.04	0	NA	1792.0
06-01-08-08-01	West Sacramento (FRH)	52	79	26955	36	54.5	51.8	2.7	10.52	147.5	3.21	0.03	0	NA	1805.0
06-01-08-08-02	West Sacramento (FRH)	52	78	25677	24	64.0	54.0	10.0	10.45	144.4	0.42	0.03	0	NA	1810.0
06-01-08-08-03	West Sacramento (FRH)	52	78	25373	25	64.0	54.0	10.0	10.45	144.4	0.42	0.03	0	NA	1816.0